

#12



OIPE

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/10/004,219B

TIME: 15:55:40

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw

3 <110> APPLICANT: Macrozyme
 4 Aerts, Johannes M.F.G.
 5 Boot, Rolf G.
 7 <120> TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
 8 its use in therapy or prophylaxis against diseases in
 9 which mucus is involved or infection diseases
 11 <130> FILE REFERENCE: 2183-5136US
 13 <140> CURRENT APPLICATION NUMBER: 10/004,219B
 14 <141> CURRENT FILING DATE: 2001-11-02
 16 <160> NUMBER OF SEQ ID NOS: 14
 18 <170> SOFTWARE: PatentIn Ver. 2.1
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 476
 22 <212> TYPE: PRT
 23 <213> ORGANISM: Artificial Sequence
 25 <220> FEATURE:
 26 <223> OTHER INFORMATION: Description of Artificial Sequence: human AMCase
 27 amino acid sequence deduced from cDNA sequence
 29 <400> SEQUENCE: 1
 30 Met Thr Lys Leu Ile Leu Leu Thr Gly Leu Val Leu Ile Leu Asn Leu
 31 1 5 10 15
 33 Gln Leu Gly Ser Ala Tyr Gln Leu Thr Cys Tyr Phe Thr Asn Trp Ala
 34 20 25 30
 36 Gln Tyr Arg Pro Gly Leu Gly Arg Phe Met Pro Asp Asn Ile Asp Pro
 37 35 40 45
 39 Cys Leu Cys Thr His Leu Ile Tyr Ala Phe Ala Gly Arg Gln Asn Asn
 40 50 55 60
 42 Glu Ile Thr Thr Ile Glu Trp Asn Asp Val Thr Leu Tyr Gln Ala Phe
 43 65 70 75 80
 45 Asn Gly Leu Lys Asn Lys Asn Ser Gln Leu Lys Thr Leu Leu Ala Ile
 46 85 90 95
 48 Gly Gly Trp Asn Phe Gly Thr Ala Pro Phe Thr Ala Met Val Ser Thr
 49 100 105 110
 51 Pro Glu Asn Arg Gln Thr Phe Ile Thr Ser Val Ile Lys Phe Leu Arg
 52 115 120 125
 54 Gln Tyr Glu Phe Asp Gly Leu Asp Phe Asp Trp Glu Tyr Pro Gly Ser
 55 130 135 140
 57 Arg Gly Ser Pro Pro Gln Asp Lys His Leu Phe Thr Val Leu Val Gln
 58 145 150 155 160
 60 Glu Met Arg Glu Ala Phe Glu Gln Glu Ala Lys Gln Ile Asn Lys Pro
 61 165 170 175
 63 Arg Leu Met Val Thr Ala Ala Val Ala Ala Gly Ile Ser Asn Ile Gln
 64 180 185 190

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/004,219B

DATE: 10/24/2002

TIME: 15:55:40

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw

```

66 Ser Gly Tyr Glu Ile Pro Gln Leu Ser Gln Tyr Leu Asp Tyr Ile His
67      195      200      205
69 Val Met Thr Tyr Asp Leu His Gly Ser Trp Glu Gly Tyr Thr Gly Glu
70      210      215      220
72 Asn Ser Pro Leu Tyr Lys Tyr Pro Thr Asp Thr Gly Ser Asn Ala Tyr
73 225      230      235      240
75 Leu Asn Val Asp Tyr Val Met Asn Tyr Trp Lys Asp Asn Gly Ala Pro
76      245      250      255
78 Ala Glu Lys Leu Ile Val Gly Phe Pro Thr Tyr Gly His Asn Phe Ile
79      260      265      270
81 Leu Ser Asn Pro Ser Asn Thr Gly Ile Gly Ala Pro Thr Ser Gly Ala
82      275      280      285
84 Gly Pro Ala Gly Pro Tyr Ala Lys Glu Ser Gly Ile Trp Ala Tyr Tyr
85      290      295      300
87 Glu Ile Cys Thr Phe Leu Lys Asn Gly Ala Thr Gln Gly Trp Asp Ala
88 305      310      315      320
90 Pro Gln Glu Val Pro Tyr Ala Tyr Gln Gly Asn Val Trp Val Gly Tyr
91      325      330      335
93 Asp Asn Ile Lys Ser Phe Asp Ile Lys Ala Gln Trp Leu Lys His Asn
94      340      345      350
96 Lys Phe Gly Gly Ala Met Val Trp Ala Ile Asp Leu Asp Asp Phe Thr
97      355      360      365
99 Gly Thr Phe Cys Asn Gln Gly Lys Phe Pro Leu Ile Ser Thr Leu Lys
100     370     375     380
102 Lys Ala Leu Gly Leu Gln Ser Ala Ser Cys Thr Ala Pro Ala Gln Pro
103 385     390     395     400
105 Ile Glu Pro Ile Thr Ala Ala Pro Ser Gly Ser Gly Asn Gly Ser Gly
106     405     410     415
108 Ser Ser Ser Ser Gly Gly Ser Ser Gly Gly Ser Gly Phe Cys Ala Val
109     420     425     430
111 Arg Ala Asn Gly Leu Tyr Pro Val Ala Asn Asn Arg Asn Ala Phe Trp
112     435     440     445
114 His Cys Val Asn Gly Val Thr Tyr Gln Gln Asn Cys Gln Ala Gly Leu
115     450     455     460
117 Val Phe Asp Thr Ser Cys Asp Cys Cys Asn Trp Ala
118 465     470     475

```

121 <210> SEQ ID NO: 2

122 <211> LENGTH: 1625

123 <212> TYPE: DNA

124 <213> ORGANISM: Artificial Sequence

126 <220> FEATURE:

127 <223> OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human

128 AMCase cDNA sequence and deduced amino acid

129 sequence

131 <220> FEATURE:

132 <223> OTHER INFORMATION: Description of Artificial Sequence: human AMCase

133 cDNA sequence and deduced amino acid sequence

135 <220> FEATURE:

136 <221> NAME/KEY: CDS

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/10/004,219B

TIME: 15:55:40

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw

```

137 <222> LOCATION: (104)..(1531)
139 <400> SEQUENCE: 2
140 gctttccagt ctggtggtga atcctccata gtctgaagcc tttgtgataa ccacagaatc 60
142 agaacatata aaaagctctg cgggactggt gctgactgca acc atg aca aag ctt 115
143                                     Met Thr Lys Leu
144                                     1
146 att ctc ctc aca ggt ctt gtc ctt ata ctg aat ttg cag ctc ggc tct 163
147 Ile Leu Leu Thr Gly Leu Val Leu Ile Leu Asn Leu Gln Leu Gly Ser
148 5 10 15 20
150 gcc tac cag ctg aca tgc tac ttc acc aac tgg gcc cag tac cgg cca 211
151 Ala Tyr Gln Leu Thr Cys Tyr Phe Thr Asn Trp Ala Gln Tyr Arg Pro
152 25 30 35
154 ggc ctg ggg cgc ttc atg cct gac aac atc gac ccc tgc ctc tgt acc 259
155 Gly Leu Gly Arg Phe Met Pro Asp Asn Ile Asp Pro Cys Leu Cys Thr
156 40 45 50
158 cac ctg atc tac gcc ttt gct ggg agg cag aac aac gag atc acc acc 307
159 His Leu Ile Tyr Ala Phe Ala Gly Arg Gln Asn Asn Glu Ile Thr Thr
160 55 60 65
162 atc gaa tgg aac gat gtg act ctc tac caa gct ttc aat ggc ctg aaa 355
163 Ile Glu Trp Asn Asp Val Thr Leu Tyr Gln Ala Phe Asn Gly Leu Lys
164 70 75 80
166 aat aag aac agc cag ctg aaa act ctc ctg gcc att gga ggc tgg aac 403
167 Asn Lys Asn Ser Gln Leu Lys Thr Leu Leu Ala Ile Gly Gly Trp Asn
168 85 90 95 100
170 ttc ggg act gcc cct ttc act gcc atg gtt tct act cct gag aac cgc 451
171 Phe Gly Thr Ala Pro Phe Thr Ala Met Val Ser Thr Pro Glu Asn Arg
172 105 110 115
174 cag act ttc atc acc tca gtc atc aaa ttc ctg cgc cag tat gag ttt 499
175 Gln Thr Phe Thr Thr Ser Val Ile Lys Phe Leu Arg Gln Tyr Glu Phe
176 120 125 130
178 gac ggg ctg gac ttt gac tgg gag tac cct ggc tct cgt ggg agc cct 547
179 Asp Gly Leu Asp Phe Asp Trp Glu Tyr Pro Gly Ser Arg Gly Ser Pro
180 135 140 145
182 cct cag gac aag cat ctc ttc act gtc ctg gtg cag gaa atg cgt gaa 595
183 Pro Gln Asp Lys His Leu Phe Thr Val Leu Val Gln Glu Met Arg Glu
184 150 155 160
186 gct ttt gag cag gag gcc aag cag atc aac aag ccc agg ctg atg gtc 643
187 Ala Phe Glu Gln Glu Ala Lys Gln Ile Asn Lys Pro Arg Leu Met Val
188 165 170 175 180
190 act gct gca gta gct gct ggc atc tcc aat atc cag tct ggc tat gag 691
191 Thr Ala Ala Val Ala Ala Gly Ile Ser Asn Ile Gln Ser Gly Tyr Glu
192 185 190 195
194 atc ccc caa ctg tca cag tac ctg gac tac atc cat gtc atg acc tac 739
195 Ile Pro Gln Leu Ser Gln Tyr Leu Asp Tyr Ile His Val Met Thr Tyr
196 200 205 210
198 gac ctc cat ggc tcc tgg gag ggc tac act gga gag aac agc ccc ctc 787
199 Asp Leu His Gly Ser Trp Glu Gly Tyr Thr Gly Glu Asn Ser Pro Leu
200 215 220 225
202 tac aaa tac ccg act gac acc ggc agc aac gcc tac ctc aat gtg gat 835

```

RAW SEQUENCE LISTING

DATE: 10/24/2002

PATENT APPLICATION: US/10/004,219B

TIME: 15:55:40

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw

```

203 Tyr Lys Tyr Pro Thr Asp Thr Gly Ser Asn Ala Tyr Leu Asn Val Asp
204      230      235      240
206 tat gtc atg aac tac tgg aag gac aat gga gca cca gct gag aag ctc 883
207 Tyr Val Met Asn Tyr Trp Lys Asp Asn Gly Ala Pro Ala Glu Lys Leu
208 245      250      255      260
210 atc gtt gga ttc cct acc tat gga cac aac ttc atc ctg agc aac ccc 931
211 Ile Val Gly Phe Pro Thr Tyr Gly His Asn Phe Ile Leu Ser Asn Pro
212      265      270      275
214 tcc aac act gga att ggt gcc ccc acc tct ggt gct ggt cct gct ggg 979
215 Ser Asn Thr Gly Ile Gly Ala Pro Thr Ser Gly Ala Gly Pro Ala Gly
216      280      285      290
218 ccc tat gcc aag gag tct ggg atc tgg gct tac tac gag atc tgt acc 1027
219 Pro Tyr Ala Lys Glu Ser Gly Ile Trp Ala Tyr Tyr Glu Ile Cys Thr
220      295      300      305
222 ttc ctg aaa aat gga gcc act cag gga tgg gat gcc cct cag gaa gtg 1075
223 Phe Leu Lys Asn Gly Ala Thr Gln Gly Trp Asp Ala Pro Gln Glu Val
224      310      315      320
226 cct tat gcc tat cag ggc aat gtg tgg gtt ggc tat gac aac atc aag 1123
227 Pro Tyr Ala Tyr Gln Gly Asn Val Trp Val Gly Tyr Asp Asn Ile Lys
228 325      330      335      340
230 agc ttc gat att aag gct caa tgg ctt aag cac aac aaa ttt gga ggc 1171
231 Ser Phe Asp Ile Lys Ala Gln Trp Leu Lys His Asn Lys Phe Gly Gly
232      345      350      355
234 gcc atg gtc tgg gcc att gat ctg gat gac ttc act ggc act ttc tgc 1219
235 Ala Met Val Trp Ala Ile Asp Leu Asp Phe Thr Gly Thr Phe Cys
236      360      365      370
238 aac cag ggc aag ttt ccc cta atc tcc acc ctg aag aag gcc ctc ggc 1267
239 Asn Gln Gly Lys Phe Pro Leu Ile Ser Thr Leu Lys Lys Ala Leu Gly
240      375      380      385
242 ctg cag agt gca agt tgc acg gct cca gct cag ccc att gag cca ata 1315
243 Leu Gln Ser Ala Ser Cys Thr Ala Pro Ala Gln Pro Ile Glu Pro Ile
244      390      395      400
246 act gct gct ccc agt ggc agc ggg aac ggg agc ggg agt agc agc tct 1363
247 Thr Ala Ala Pro Ser Gly Ser Gly Asn Gly Ser Gly Ser Ser Ser Ser
248 405      410      415      420
250 gga ggc agc tcg gga ggc agt gga ttc tgt gct gtc aga gcc aac ggc 1411
251 Gly Gly Ser Ser Gly Gly Ser Gly Phe Cys Ala Val Arg Ala Asn Gly
252      425      430      435
254 ctc tac ccc gtg gca aat aac aga aat gcc ttc tgg cac tgc gtg aat 1459
255 Leu Tyr Pro Val Ala Asn Asn Arg Asn Ala Phe Trp His Cys Val Asn
256      440      445      450
258 gga gtc acg tac cag cag aac tgc cag gcc ggg ctt gtc ttc gac acc 1507
259 Gly Val Thr Tyr Gln Gln Asn Cys Gln Ala Gly Leu Val Phe Asp Thr
260      455      460      465
262 agc tgt gat tgc tgc aac tgg gca taaacctgac ctggtctata ttccctagag 1561
263 Ser Cys Asp Cys Cys Asn Trp Ala
264      470      475
266 ttccagtctc ttttgcttag gacatgttgc ccctacctaa agtcttgcaa taaaatcagc 1621
268 agtc 1625

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/004,219B

DATE: 10/24/2002

TIME: 15:55:40

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw

```

271 <210> SEQ ID NO: 3
272 <211> LENGTH: 1525
273 <212> TYPE: DNA
274 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse
278     AMCase cDNA sequence and deduced amino acid
279     sequence
281 <220> FEATURE:
282 <223> OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
283     cDNA sequence and deduced amino acid sequence
285 <220> FEATURE:
286 <221> NAME/KEY: CDS
287 <222> LOCATION: (1)..(1419)
289 <400> SEQUENCE: 3
290 atg gcc aag cta ctt ctc gtc aca ggt ctg gct ctt ctg ctg aat gct      48
291 Met Ala Lys Leu Leu Leu Val Thr Gly Leu Ala Leu Leu Leu Asn Ala
292   1             5             10             15
294 cag ctg ggg tct gcc tac aat ctg ata tgc tat ttc acc aac tgg gcc      96
295 Gln Leu Gly Ser Ala Tyr Asn Leu Ile Cys Tyr Phe Thr Asn Trp Ala
296           20           25           30
298 cag tat cgg cca ggt ctg ggg agc ttc aag cct gat gac att aac ccc      144
299 Gln Tyr Arg Pro Gly Leu Gly Ser Phe Lys Pro Asp Asp Ile Asn Pro
300       35       40       45
302 tgc ctg tgt act cac ctg atc tat gcc ttt gct ggg atg cag aac aat      192
303 Cys Leu Cys Thr His Leu Ile Tyr Ala Phe Ala Gly Met Gln Asn Asn
304   50           55           60
306 gag atc acc acc ata gaa tgg aat gat gtt act ctc tat aaa gct ttc      240
307 Glu Ile Thr Thr Ile Glu Trp Asn Asp Val Thr Leu Tyr Lys Ala Phe
308  65           70           75           80
310 aat gac ttg aaa aac agg aac agc aaa ctg aaa acc ctc ctg gca att      288
311 Asn Asp Leu Lys Asn Arg Asn Ser Lys Leu Lys Thr Leu Leu Ala Ile
312           85           90           95
314 gga ggc tgg aac ttt gga act gct cct ttc act acc atg gtt tcc act      336
315 Gly Gly Trp Asn Phe Gly Thr Ala Pro Phe Thr Thr Met Val Ser Thr
316       100       105       110
318 tct cag aac cgc cag acc ttc att acc tca gtc atc aaa ttt ctg cgt      384
319 Ser Gln Asn Arg Gln Thr Phe Ile Thr Ser Val Ile Lys Phe Leu Arg
320       115       120       125
322 cag tat ggg ttt gat gga ctg gac ctg gac tgg gaa tac cca ggc tca      432
323 Gln Tyr Gly Phe Asp Gly Leu Asp Leu Asp Trp Glu Tyr Pro Gly Ser
324   130       135       140
326 cgt ggg agc cct cct cag gac aag cat ctc ttc act gtc ctg gtg aag      480
327 Arg Gly Ser Pro Pro Gln Asp Lys His Leu Phe Thr Val Leu Val Lys
328 145       150       155       160
330 gaa atg cgt gaa gct ttt gag cag gag gct att gag agc aac agg ccc      528
331 Glu Met Arg Glu Ala Phe Glu Gln Glu Ala Ile Glu Ser Asn Arg Pro
332       165       170       175
334 aga ctg atg gtt act gct gct gta gct ggt ggg att tcc aac atc cag      576

```

VERIFICATION SUMMARY

DATE: 10/24/2002

PATENT APPLICATION: US/10/004,219B

TIME: 15:55:41

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\10242002\J004219B.raw